SYSTEM AND METHOD FOR ENHANCING THE INTELLIGIBILITY OF RECEIVED SPEECH IN A NOISE ENVIRONMENT

RELATED APPLICATIONS

This application claims the benefit of the filing date of United States provisional patent application serial no. 60/207,273, filed on May 30, 2000.

FIELD OF THE INVENTION

The present invention relates generally to audio communication devices, and specifically to intelligibility enhancement of speech received by a communications device.

BACKGROUND

Receiving and understanding a voice communication received using a communications device while in a noise environment is problematic. Ambient or background noise near the receiver, the near-end, may at times approach or exceed the volume levels of the received speech causing a listener to be unable to understand the message.

Simply increasing the volume of the received speech is not very effective in increasing the intelligibility of the speech. Speech amplitude levels fluctuate throughout a sentence, from word to word, and even from syllable to syllable. Increasing the volume to compensate for a low amplitude portion of the speech may cause the receiver to clip when a high amplitude portion of the speech is received. Adjusting the volume to compensate for a high amplitude portion may render a subsequent portion having a low amplitude unintelligible.

U.S. Pat. No. 5,907,823 describes a dynamic range compression method which based on an instantaneous measurement of a received speech level or amplitude and an estimate of the ambient and far-end noise levels adjusts the received signal to enhance intelligibility. Fig. 1 shows a system according to the '823 patent. The dynamic range compression module

1

receives (1) the received speech signal, (2) a signal indicating an estimate of the far-end noise level, and (3) a signal indicating an estimate of the near-end or ambient noise level. The dynamic range compression module performs a transform of the received speech according to the transform pattern shown in Fig. 2.

Several drawback in the prior art systems exist which poorly address the issue of rapid variations in speech volume and very low speech levels.

BRIEF DESCRIPTION OF THE DRAWINGS

The subject matter regarded as the invention is particularly pointed out and distinctly claimed in the concluding portion of the specification. The invention, however, both as to organization and method of operation, together with objects, features, and advantages thereof, may best be understood by reference to the following detailed description when read with the accompanying drawings in which:

FIG. 1 is a block diagram showing a speech intelligibility enhancement system according to the prior art;

FIG. 2 is a graph showing the transform function or pattern of the system in Fig. 1;

FIG. 3 is a block diagram showing a speech intelligibility enhancement system according to the present invention; and

FIG 4 is a flow diagram showing the steps performed by a speech intelligibility enhancement system according to the present invention.

It will be appreciated that for simplicity and clarity of illustration, elements shown in the figures have not necessarily been drawn to scale. For example, the dimensions of some of the elements may be exaggerated relative to other elements for clarity. Further, where considered appropriate, reference numerals may be repeated among the figures to indicate corresponding or analogous elements.

2

DETAILED DESCRIPTION

In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the invention. However, it will be understood by those skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known methods, procedures, components and circuits have not been described in detail so as not to obscure the present invention.

Unless specifically stated otherwise, as apparent from the following discussions, it is appreciated that throughout the specification discussions utilizing terms such as "processing", "computing", "calculating", "determining", or the like, refer to the action and/or processes of a computer or computing system, or similar electronic computing device, that manipulate and/or transform data represented as physical, such as electronic, quantities within the computing system's registers and/or memories into other data similarly represented as physical quantities within the computing system's memories, registers or other such information storage, transmission or display devices.

Embodiments of the present invention may include apparatuses for performing the operations herein. This apparatus may be specially constructed for the desired purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs) electrically programmable read-only memories (EPROMs), electrically erasable and programmable read only memories (EEPROMs),

magnetic or optical cards, or any other type of media suitable for storing electronic instructions, and capable of being coupled to a computer system bus.

The processes and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the desired method. The desired structure for a variety of these systems will appear from the description below. In addition, embodiments of the present invention are not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the inventions as described herein.

As part of the present invention, a received signal containing speech may be divided into frames having some predefined number of samples. A frame may then be analyzed to determine an instantaneous peak value, and from the instantaneous peak values of two or more consecutive frames, a long term peak value may be determined for each frame. For a given frame, an automatic gain control ("AGC") value may be calculated in relation to the difference between a desired volume or amplitude level and the frame's long term peak value. A noise factor between 0 and 1 may be factored into the AGC value. The samples within a frame may be increased or adjusted by the AGC value.

A dynamic range compression ("DRC") gain value may be calculated for each frame and the samples within the frame may be adjusted by the RDC gain value. A noise factor may also me factored into the DRC gain value. The DRC gain value for a frame may be calculated as the difference between the frame's long term peak value and its instantaneous peak value, multiplied by a maximum gain value divided by ΔL, where ΔL is a range value.

4

A target gain value for a frame, which is the sum of the frame's AGC and RDC gains, may be smoothed from frame to subsequent frame. One method of smoothing is to average a frame's calculated target gain with the previous frame's target gain value.

Turing now to Fig. 3, there is shown a system according to the present invention. A received signal RX is decoded by decoder 10 and speech within the signal is detected by a voice activity detector 20. A noise suppressor 30 may remove noise from the signal and pass the signal to the AGC unit 40. The AGC unit 40 may compute an AGC gain value and apply the value to the signal, and the DRC unit 50 may also compute and apply a DRG gain value to the signal. Both the AGC 40 and the DRC 50 may receive a signal from a TX noise estimator 60 indicating an estimate of the noise on the receiver or near side.

Fig. 4 is a flow diagram showing the steps performed by a speech intelligibility enhancement system according to the present invention. For each frame of a received signal an instantaneous peak amplitude or value is computed as part of step 100. The size of a frame may vary (e.g. 0.1 to 100 ms). The instantaneous peak amplitude is simple the highest absolute value of any sample in the frame in dB:

$$\text{instantaneous peak amplitude} = 20*\log(\max(\text{abs}(x[n])))$$

As part of step 200, a long term peak amplitude value is calculated for a frame having speech present therein. Step 300 shows that speech presence in a frame may be indicated by a flag determined either by a voice activity detection algorithm for a fixed-rate vocoder, or by the vocoder rate being maximal for a variable rate vocoder. A frame's long term peak amplitude value may be determined using the following logic:

```
If(voice present)
    If(instantaneous peak value > long term peak amplitude value)
        long term peak amplitude value = instantaneous peak value
    Else if(instantaneous peak value>Minimum Acceptable Value)
        long term peak amplitude value = long term peak amplitude value –
decrement peak value
```

5

The decrement peak value is chosen to decrease the long term peak amplitude value by a nominal amount.

As part of step 400 an AGC gain value is calculated. The AGC gain may be calculated by subtracting from a target or desired level value L1 the frame's long term peak amplitude value. In one embodiment of the present invention the calculated AGC gain may also be multiplied by a TX noise factor calculated as part of steps 500, 600 and 700.

A DRC gain value is computed as part of step 800. The DRC gain may be computed as the difference between a frame's long term peak value and its instantaneous peak value, multiplied by a maximum gain value divided by ΔL:

DRC gain value=(long term peak value-instantaneous peak value)×Max DRC Gain/Δ L,

Where ΔL is a range value typically 20dB and Max RDC gain is also typically 20dB. The DRC gain value may also be multiplied by a noise factor such as the one calculated in steps 500, 600 and 700.

A target gain value may be calculated in step 900. The target gain value may be the sum of the AGC and RDC gain values. Smoothing of the computed target gain is performed in step 1000. Each sample in each frame may be adjusted by the smoothed target gain value.

While certain features of the invention have been illustrated and described herein, many modifications, substitutions, changes, and equivalents will now occur to those skilled in the art. It is, therefore, to be understood that the appended claims are intended to cover all such modifications and changes as fall within the true spirit of the invention. **A description of a second embodiment is attached as Appendix A.**

Appendix A

RX-AVC: Automatic Volume Control Algorithm
(Internal distribution only)

1   Background

The RX-AVC is an advanced audio function that automatically controls the received speech level and dynamic range. The AVC is a collection of 3 functions. The 3 functions can be activated both separately and together. The 3 functions are:

1. Automatic gain control (AGC)

2. Automatic volume increase and dynamic range compression (DRC) as a function of TX noise level

3. Dynamic range compression (DRC) for speakerphone

The following block diagram shows how the RX-AVC fits within the other speech modules in the RX and TX paths.
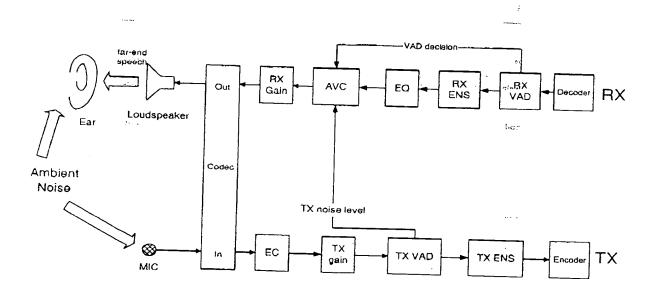

Figure 1: Audio path with RX-AVC.

2 Algorithm description

The total AVC gain is the sum of 3 gains:

1. AGC

2. DRC gain weighted by a maximum between two numbers: (1) a weight proportional to the TX noise level, (2) DRC GAIN FACTOR.

3. Additional constant gain whose value is weighted by TX noise level

Enabling flags (AGC_FLAG and NOISE_RESPONSE_FLAGS) gate the AGC and the noise-weighted gains. Figure 2 shows the algorithm block diagram.
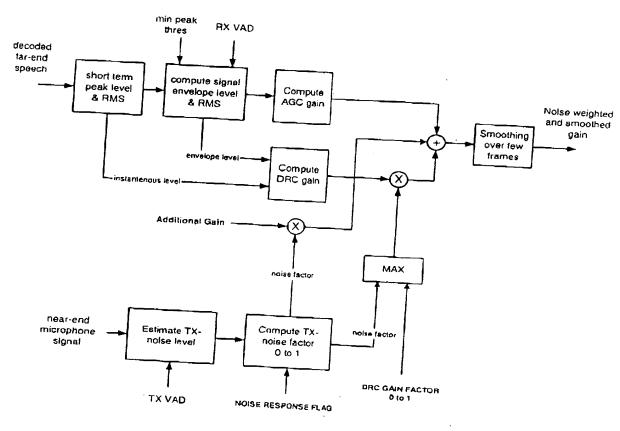

Figure 2: block diagram

2.1 Blocks description

2.1.1 short-term peak & short-term rms level

Computes maximal sample energy over a sub frame several samples (default=16)
Computes rms energy over the sub frame

2.1.2 Compute signal envelope level & speech-RMS

Envelope is updated if RX-VAD detected speech and if the local peak is above a threshold:

```
if (RX-VAD && local_peak > MIN_ENVELOPE)
{
    if (local_peak > envelope)
        envelope = min(local_peak, envelope + INCREMENT_PEAK) ;
    else
        envelope -= DECREMENT_PEAK ;

    if (rms > long_term_rms)
        long_term_rms += INCREMENT_RMS ;
    else
        long_term_rms -= INCREMENT_RMS ;
}
```

2.1.3 Compute AGC gain

Basic relation: $agc\_gain = (LEVEL1 - envelope)$;
Limitation 1: $agc\_gain = min(MAX\_AGC\_GAIN, agc\_gain)$;
Limitation2: $agc\_gain = min(MAX\_RMS - long\_term\_rms, agc\_gain)$;

2.1.4 Compute DRC gain

Given the following noise dependent and RX- signal independent parameter:
$drc\_gain = MAX\_DRC\_GAIN * max(drc\_gain\_factor, noise\_factor)$;

Compute for each sub-frame the following gain (frame_drc_gain) as a function of its local peak and the envelope level as depicted in the following figure:
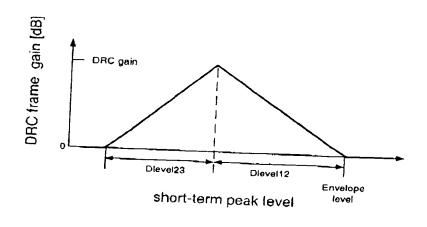
Figure 3: DRC gain function

2.1.5 Estimate TX noise level

See Period_alg.doc

2.1.6 Compute Noise factor
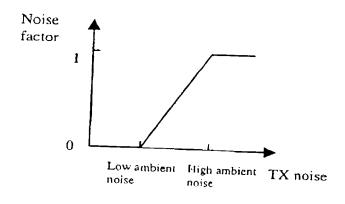
Noise factor is a function of noise level:

Figure 4: Noise factor

4

2.1.7 Gain smoothing
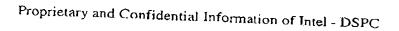
Signed exponential smoothing:

```
if (target_gain > smooth_gain_db)
    alfa = ALFA  UP ;
  else
    alfa = ALFA_DOWN ;


  smooth_gain_db = smooth_gain_db * (1-alfa) + alfa * target_gain ;
```

3 Parameters

3.1 Tunable parameters

AGC_FLAG

AGC enable/disable control bit

NOISE_RESPONSE_FLAG

Enable/disable control bit for the ambient noise-induced automatic volume increase.

DRC_GAIN_FACTOR

Controls the amount of DRC, ranging from 0 to 0x7FFF. Defaul = 0.

LEVEL1

Target level for the RX-signal envelope level. Default value: -3 dBov (3 dB below clipping level).

HIGH/LOW_AMBIENT_NOISE:

See figure 3. The default values are -39 dBov (low) and -21 dBov (High). Increasing these thresholds will reduce the effect of the RX-AGC to noise. The thresholds should be modified if the TX analog path gain is changed. For example, if the TX analog codec gain is increased by X dB, these thresholds should be increased by the same X dB.

MAX_AGC_GAIN

Upper limit on the AGC gain in the module "Compute AGC gain" (see 2.1.3). The default value is set to 21 dB.
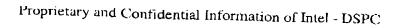
MAX_DRC_GAIN

Upper limit on the DRC gain in the module "compute DRC gain" (see 2.1.4). The default value is set to its maximum allowed value of 21 dB. Decreasing this parameter will decrease the effect of the DRC.

MAX_RMS

The AGC gain is limited so that the signal RMS after amplification does not exceed MAX_RMS (default = -15 dBov).

MIN_ENVELOPE

Signals whose envelope is smaller than this threshold do not affect the AGC.

brief

3.2 Soft constants (default)

NUM_SUB_FRAMES	(10)
SUB_FRAM_LEN	(FRAM_LEN/NUM_SUB_FRAMES = 16)
LOG_SUB_FRAM_LEN	$(10*\log10(SUB\_FRAM\_LEN) = 12)$
DLEVEL12 (see Figure 3)	(24)
DLEVEL23 (see Figure 3)	(26)
DECREMENT_PEAK	(3*SUB_FRAM_LEN/8000 = 3 dB/sec)
INCREMENT_PEAK	(1 = 1dB/sub_frame)
INCREMENT_RMS	(10*SUB_FRAM_LEN/8000 = 10 dB/sec)
ALFA_UP (see 2.1.7)	0.1)
ALFA_DOWN (see 2.1.7)	(0.5)